



Facultad de Ciencias de la Administración

**Escuela de Ingeniería en Ciencias de la
Computación**

**EVALUACIÓN PARA DISMINUIR LA
PÉRDIDA EN LA TRANSCRIPCIÓN DE
AUDIOS DE UN CENTRO DE COMANDO Y
CONTROL**

**Trabajo de titulación previo a la obtención del
grado de Ingeniero en Ciencias de la
Computación**

Autor:

Luis Daniel Márquez Castillo

Director:

Ing. Marcos Patricio Orellana Cordero

Cuenca – Ecuador

Año

2024

DEDICATORIA

A mis padres, mis hermanos y a los amigos de siempre, por su apoyo incondicional y por ser mi fuerza en cada paso de este camino. A todos ustedes, con gratitud infinita, dedico este trabajo.

AGRADECIMIENTO

Al Laboratorio de Investigación y Desarrollo Informático de la Universidad del Azuay, a la Universidad y a los docentes que me formaron y apoyaron durante este tiempo, por el respaldo y la guía que hicieron posible la realización de este trabajo.

Índice de Contenidos

DEDICATORIA	i
AGRADECIMIENTO	ii
Índice de Contenidos	iii
Índice de tablas	iv
Índice de figuras	v
RESUMEN	vi
ABSTRACT	vi
EVALUACIÓN PARA REDUCIR LA PÉRDIDA DE TRANSCRIPCIÓN DE AUDIO EN UN CENTRO DE CONTROL	1
1. Introducción	1
1.1 Objetivos	1
1.1.1 Objetivo general	1
1.1.2 Objetivos específicos	1
1.2 Marco teórico	1
1.3. Revisión de literatura	3
2. Métodos	5
2.1 Comprensión del negocio.....	5
2.2 Comprensión de los datos	6
2.2.1 Descripción general del conjunto de datos.....	6
2.2.2 Particularidades lingüísticas.....	6
2.2.3 Calidad del audio	7
2.3 Preparación de los datos.....	7
2.3.1 Descripción de los datos	7
2.3.2 Eliminación de datos atípicos	7
2.3.3 Selección de la muestra.....	10
2.4 Modelado	11
2.5 Evaluación.....	12
3. Resultados	13
4. Discusión	15
5. Conclusiones	17
6. Referencias	18

Índice de tablas

Tabla 1	Tabla de Promedios de Frecuencia de Palabras según Severidad.....	8
Tabla 2	Tiempo de transcripción de audio.....	14
Tabla 3	Promedios del WER y Desviación estándar de los resultados en el conjunto de datos de prueba.....	14
Tabla 4	Promedio de mejora de los modelos propios, respecto al Whisper pre-entrenado..	14
Tabla 5	Tabla estadística de los valores WER de modelo Large	15

Índice de figuras

Figura 1 Meta-modelo SPEM 2.0, adaptación pasos 2,3,4 y 5 de la metodología CRISP-DM	6
Figura 2 Análisis estadístico de normalidad y desviación estándar del número de palabras por audio.....	9
Figura 3 Análisis estadístico de normalidad y desviación estándar del número de palabras por audio de la alerta Roja.....	9
Figura 4 Análisis estadístico de normalidad y desviación estándar del número de palabras por audio de la alerta No Roja.....	10
Figura 5 Comparación entre las estadísticas de la tasa WER obtenida en cada versión del modelo Large	15
Figura 6 Comparación del promedio del porcentaje del promedio WER entre el modelo pre-entrenado y los ajustados	17

EVALUACIÓN PARA DISMINUIR LA PÉRDIDA EN LA TRANSCRIPCIÓN DE AUDIOS DE UN CENTRO DE COMANDO Y CONTROL

RESUMEN

Los avances en el procesamiento de lenguaje natural (NLP) y en los modelos de reconocimiento de voz (ASR) han mejorado la transcripción automática, pero siguen enfrentando limitaciones en entornos desafiantes como los centros de comando y control, donde la calidad del audio y las particularidades del lenguaje afectan a la precisión. Este estudio se enfoca en reducir la tasa de error de palabras (WER) en las transcripciones de audios del Servicio Integrado de Seguridad ECU 911, provenientes de la zona austral de Ecuador, principalmente de la ciudad de Cuenca. Se evaluaron tres escenarios: la versión original del modelo, el modelo con ajuste-fino y el modelo con ajuste-fino acompañado de un filtro paso-bajo. Dependiendo del contexto de ruido y del tamaño del modelo, se lograron mejoras en la precisión de las transcripciones, con reducciones de WER entre un 1% y un 22%. Estos resultados demuestran que la combinación de ajuste-fino y tratamiento de señales es efectiva para mejorar la calidad de las transcripciones, especialmente en ambientes con ruido y características acústicas complejas.

Palabras clave: Inteligencia artificial, Procesamiento de la información, Procesamiento del lenguaje natural, Reconocimiento de la voz, Sistemas de información

EVALUATION TO REDUCE AUDIO TRANSCRIPTION LOSS IN A COMMAND-AND-CONTROL CENTER ABSTRACT

Advances in natural language processing (NLP) and automatic speech recognition (ASR) models have enhanced automatic transcription, yet they continue to face limitations in challenging environments, such as command-and-control centers, where audio quality and language characteristics impact accuracy. This study focuses on reducing the word error rate (WER) in transcriptions of audio from the Integrated Security Service ECU 911, originating from the southern region of Ecuador, primarily from the city of Cuenca. Three scenarios were evaluated: the original model, the fine-tuned model, and the fine-tuned model with a low-pass filter. Depending on the noise context and model size, transcription accuracy improved, achieving WER reductions between 1% and 22%. These findings demonstrate that combining fine-tuning with signal processing is effective in enhancing transcription quality, particularly in noisy environments and complex acoustic conditions.

Keywords: Artificial intelligence, Information processing, Information systems, Natural language processing, Speech recognition